

# PRIVACY RESEARCH IN CENSUS 2000

Quality assurance procedures were applied throughout the creation of this report.

This topic report integrates findings and provides context and background for interpretation of results from Census 2000 evaluations, tests, and other research undertaken by the U.S. Census Bureau. It is part of a broad program, the Census 2000 Testing, Experimentation, and Evaluation program, designed to assess Census 2000 and to inform 2010 Census planning.

Prepared by Eleanor Singer  
University of Michigan

---

Linda Brudvig, Project Manager  
Planning, Research, and Evaluation Division

## CONTENTS

1.	BACKGROUND .....	1
2.	CHANGES IN ATTITUDES ABOUT PRIVACY, CONFIDENTIALITY AND DATA SHARING OVER TIME .....	3
2.1	Methods .....	3
2.2	Limitations .....	4
2.3	Trends in beliefs about confidentiality .....	4
2.4	Trends in attitudes toward privacy .....	7
2.5	Attitudes toward confidentiality in Puerto Rico .....	7
2.6	Trends in attitudes toward data sharing .....	7
3.	THE EFFECT OF THE CENSUS OUTREACH CAMPAIGN ON ATTITUDES TOWARD CONFIDENTIALITY AND DATA SHARING .....	9
4.	THE IMPACT OF NEGATIVE PUBLICITY ON PRIVACY CONCERNS .....	11
5.	THE EFFECT OF PRIVACY-CONFIDENTIALITY CONCERNS ON CENSUS PARTICIPATION .....	13
6.	PUBLIC PERCEPTIONS OF AGENCY CONFIDENTIALITY PRACTICES .....	15
7.	THE EFFECT OF A REQUEST FOR SOCIAL SECURITY NUMBERS .....	16
8.	THE ROLE OF THE PARTNERSHIP PROGRAM IN PRIVACY ATTITUDES ....	18
9.	ATTITUDES AND BEHAVIOR .....	18
10.	THE ROLE OF INFORMED CONSENT .....	19
11.	SUMMARY OF FINDINGS .....	20
11.1	Trends in attitudes, 1995-2000 .....	20
11.2	The effect of Census Bureau public relations activities on beliefs about confidentiality .....	21
11.3	Attitudes and behavior .....	21
11.4	Willingness to provide SSN's .....	22
12.	CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE RESEARCH .....	22
12.1	Conclusions .....	22
12.2	Recommendations .....	23
	Acknowledgments.....	25
	References.....	26

## 1. BACKGROUND

The Census 2000 Testing, Experimentation, and Evaluation Program provides measures of effectiveness for the Census 2000 design, operations, systems, and processes and provides information on the value of new or different methodologies. The results and recommendations from these analyses provide valuable information crucial to planning the 2010 Census. By providing measures of how well Census 2000 was conducted, this program fully supports the Census Bureau's strategy to integrate the 2010 planning process with ongoing Master Address File/TIGER enhancements and the American Community Survey. The purpose of the report that follows is to synthesize results from related Census 2000 evaluations, experiments, and other assessments to make recommendations for planning the 2010 Census. Census 2000 Testing, Experimentation, and Evaluation reports are available on the Census Bureau's Internet site at: <http://www.census.gov/pred/www/>.

The mail return to the 1990 U.S. census averaged 64.6 percent, some 10 percentage points less than in 1980 and 5 points less than had been anticipated by the Census Bureau (Singer, Mathiowetz, and Couper, 1993). One hypothesis put forward for the reduced return rate was increased public concern about privacy, documented in a series of surveys by the Harris Organization (Westin, 1990), and about confidentiality. In this report, concern about confidentiality refers to a desire to keep information already given to one agent out of the hands of others; concern about privacy refers to a desire to keep information out of the hands of others altogether. Although there is some evidence that the public may be blurring the distinction between these concepts (Martin, 2000), the distinction appears to be a meaningful one in much of the research reported here. Although the Outreach Evaluation Study, carried out by the Census Bureau in 1990, found that the large majority of respondents believed that census data are kept confidential (Fay, Bates, and Moore, 1991:18), and that such beliefs had not declined since the last decennial census, it also documented a significant change in the relationship between trust in the Census Bureau's assurance of confidentiality and self-reported census return rate. Whereas trust was not predictive of self-reported returns in 1980, it was predictive of such returns in 1990 (*ibid.* and table 4), with some 17 points separating the self-reported return rates of those with a high and a low degree of trust.<sup>1</sup>

In an analysis of actual census mail return rates and attitudes toward privacy and confidentiality (as measured in the Survey of Census Participation, carried out by the National Opinion Research Center in the summer of 1990), Singer, Mathiowetz, and Couper (1993) found that both attitudes were predictive of actual returns, with concerns about confidentiality, measured by a series of items all pertaining to the census, the stronger predictor of the two.<sup>2</sup> In a subsequent

---

<sup>1</sup> In 1999 and 2000, the relationship between trust in the Census Bureau's promise of confidentiality and self-reported return of the census form was smaller but still statistically significant (Singer, 2001, p. 342).

<sup>2</sup> The privacy index consisted of eight items, only two of which dealt explicitly with the Census Bureau or census.

analysis that pitted concerns about privacy and confidentiality against other attitudes, demographic characteristics, and various measures of competing demands as well as access and capacity, Couper, Singer, and Kulka (1998) demonstrated that confidentiality concerns (but not concerns about privacy) remained a significant predictor of mail returns to the 1990 census.

Adding to the Census Bureau's unease was a National Academy of Sciences panel recommendation that it consider using administrative records to improve the accuracy of the Census 2000 count (Steffey and Bradburn, 1994). It was hypothesized that such data sharing among federal agencies, if it became public knowledge, might increase existing confidentiality concerns, as might a request for the respondent's Social Security number (SSN) to facilitate the merging of information.

As a result of these various developments, the Census Bureau in the early 1990s embarked on a program of privacy-related research, including focus groups, large-scale experiments, and commitment to support a series of cross-sectional surveys that would track attitudes about privacy and confidentiality, especially as these related to the decennial census and the proposal to supplement the traditional count by use of administrative records. This report synthesizes the findings resulting from that program under the following headings: Changes in attitudes about privacy, confidentiality and data sharing over time; the effect of the Census outreach campaign on attitudes toward confidentiality and data sharing; the impact of negative publicity on privacy concerns; the effect of privacy-confidentiality concerns on census participation; public perceptions of agency confidentiality practices; the effect of a request for Social Security numbers; the role of the partnership program in privacy attitudes; attitudes and behavior; and the role of informed consent. The research projects serving as the basis for the evaluation were the Surveys of Privacy Attitudes (SPA); the Social Security Number, Privacy Attitudes, and Notification experiment (SPAN), which examined the effect of requesting Social Security numbers on unit and item nonresponse to the census form; the Survey of Partners, which questioned organizations that participated in the outreach program on what they thought was effective or ineffective about the campaign; the report of focus groups held in Puerto Rico on why households did not mail back the Census 2000 questionnaire; an ethnographic investigation of people's privacy schemas; and an Internet survey of privacy attitudes during Census 2000.

However, drawing inferences from this research for the public's behavior seven years from now is risky. The world has changed drastically since Census 2000 and the research reviewed here were carried out. A terrorist attack leveled the World Trade Center and damaged the Pentagon. The future, in terms of other terrorist attacks on the United States and more stringent security laws, is uncertain. All of these events have potential implications for the climate of public opinion that will surround the next decennial census, but it is impossible at present to predict either what that climate will be or how it will affect data collection efforts.

Since Census 2000, two pieces of legislation diametrically opposed in their implications have both been enacted into law. The E-Government Act, passed in November 2002, provided unprecedented protections for the confidentiality of data collected by government agencies. In principle, the existence of this legislation should make it easier to reassure the public that the

information they provide to the Census Bureau, as well as other government agencies, cannot be used in administrative proceedings against them, and that the confidentiality of this information is protected by strong legislation and appropriate penalties. However, the USA Patriot Act was also enacted into law since Census 2000, and extensions to it are being planned as this is being written. The implications of that legislation for the ability of the Justice Department and other law enforcement agencies to gain access to data regardless of the confidentiality assurances given by other agencies have not yet been tested. As a result, it is by no means clear what the net effect of these two pieces of legislation will be.

## **2. CHANGES IN ATTITUDES ABOUT PRIVACY, CONFIDENTIALITY, AND DATA SHARING OVER TIME<sup>3</sup>**

### **2.1 Methods**

The findings reported in this section are based largely on four surveys of the telephone population 18 and over residing in the contiguous 48 states carried out between 1995 and 2000. The first was developed in consultation with the Census Bureau as part of the University of Maryland's 1995 Joint Program in Survey Methodology practicum. The second, which used a questionnaire virtually identical to that in 1995, was carried out by Westat in 1996 under contract to the Census Bureau. The third and fourth, done in July through October 1999, just before the start of the public relations campaign and nationwide field recruiting for Census 2000, and from April to July of 2000, after delivery of census forms to U.S. households, were done by the University of Michigan under contract with the Census Bureau, with data collected by The Gallup Organization.

All four surveys used virtually identical methods and achieved very similar response rates. All were random digit dialed surveys with one member of the household aged 18 or over randomly selected after household listing by the interviewer. The response rates for each survey are shown below:

Year	Sample Size	Response Rate
1995	1443	61 <sup>a</sup>
1996	1215	60 <sup>b</sup>
1999	1677	62 <sup>b</sup>
2000	1978	61 <sup>b</sup>

<sup>a</sup> Interviews divided by the total sample less businesses, nonworking numbers, and numbers that were never answered after a minimum of 20 calls.

<sup>b</sup> Interviews divided by the total sample less businesses, nonworking numbers, and the estimated number of ineligibles among the noncontacts.

### **2.2 Limitations**

A number of limitations attend the findings summarized in Section 2. First, the population covered excludes adults living in non-telephone households and non-English speakers (except for Spanish speakers in 1999 and 2000). Second, compared to face-to-face surveys, the response

---

<sup>3</sup> This section is an abridged version of material reported in Singer (2001).

rates obtained are relatively low; third, attitudes about confidentiality and data sharing may belong to the realm of nonattitudes—that is, attitudes that are created on the spot and therefore ephemeral; finally, the organizations carrying out the surveys have changed over the years.

These limitations are discussed in more detail in Singer (2001). In brief, there is reason to believe that the bias introduced by nonresponse to the survey is relatively small, but that it serves to underestimate privacy concerns and opposition to data sharing. There is also reason to believe that respondents answer questions about novel phenomena not in random fashion, but rather in light of their attitudes toward known, related stimuli. Examination of the survey findings and related behaviors suggests that attitudes bear a nontrivial relationship to relevant behavior, especially when the object of the attitude and the behavior are identical (i.e., provision of a SSN).

Changes in the organizations carrying out the surveys are another potential threat to the data, and, given that three different survey organizations carried out four surveys, there is no way to control for this effect. However, many so-called “house” effects are attributable to variations in question order or differences in probing of Don’t Know responses. Question order was virtually identical across the four surveys discussed here, and Don’t Know and Not Sure rates are very similar across the four surveys. Furthermore, the pattern of changes in responses over time varies from question to question, making it unlikely that there are consistent house effects in these data. Changes in sample composition that might result from different calling algorithms and different refusal conversion strategies are compensated to some extent by weighting the data to known census distributions, although such weighting obviously cannot compensate for attitudinal differences that might be associated with variations in sample composition due to differential recruitment and follow-up strategies. This last limitation potentially affects the accuracy of the estimated trends from 1995 to 2000. Because the same organization carried out the 1999 and 2000 surveys, estimates of attitude stability and change during those two years should be less affected.

A potential source of bias in three of the surveys is the introduction, which in 1996, 1999, and 2000 mentioned both the fact that the survey was “a study of people’s attitudes about whether government agencies keep information about them private” and that it was being done on behalf of the U.S. Census Bureau. This may have reduced the participation of people most concerned about privacy, as well as that of people with unfavorable attitudes toward the Census Bureau.

### **2.3 Trends in beliefs about confidentiality**

Trends in beliefs about the Census Bureau’s treatment of personal information were measured in several different ways on these surveys. Early in the interview, respondents were asked for their beliefs about Census Bureau practices. Later questions probed their knowledge of the laws governing confidentiality practices, and then those knowledgeable about the relevant laws were asked whether they trusted the Census Bureau to obey them. Finally, at the very end of the interview, respondents were asked several questions about potential misuses of census data involving breaches of confidentiality. Most questions were asked in all four years, but some

were asked in only three, as a result of our experience with the 1995 survey.

The first question designed to probe beliefs about actual practices asked, “Do you believe other agencies, outside the Census Bureau, can or cannot get people’s names and addresses along with their answers to the census, or are you not sure?” The introduction to the question referred back to the demographic questions asked on the short form and informed people that “the person in the household who fills out the form must list the full name of everyone who lives there along with each person’s age, sex, race, [and marital status.]” The second question, asked for the first time in 1996 to assess whether use of the term “confidentiality” would change the pattern of responses, was, “Do you think the Census Bureau does or does not protect the confidentiality of this information, or are you not sure?” with an introduction identical to that already quoted. Respondents in 1996 were randomly assigned to one question or the other. Finally, in 1999 and 2000, in order to try to clarify earlier inconsistencies, one third of the sample was asked both of these questions (with the order of questions randomized), followed by an open-ended question about the meaning of confidentiality to the respondent.

Responses to the two questions inquiring into beliefs about Census Bureau practices show a significant increase between 1996 and 2000 in the proportion giving the correct response (that other agencies cannot get the data, and that the Census Bureau protects confidentiality)--from 6.1 percent to 17.3 percent in the case of “can get (Singer *et al.*, 2001, Table 2.7),” and from 12.9 percent to 25.1 percent in the case of confidentiality (*ibid.*, Table 2.8). Unlike later questions discussed in this section, these questions offered an explicit Not Sure category to respondents. The very large proportion of Not Sure answers, which is perhaps the most striking feature of both tables, is, therefore, a function both of the public’s lack of information and of the response options offered by the question; cf. Schuman and Presser (1981). In 1996, for example, when the questions were asked both with an explicit Not Sure option and, in split-ballot form, without such an option, the Not Sure rate shown dropped from 46.8 percent to 7.7 percent; however, the ratio of correct to incorrect responses did not change (Kerwin and Edwards, 1996, Table 7).<sup>4</sup>

Data comparable to those reported above are also available from a National Research Council (1979) study inquiring into confidentiality concerns as factors in survey response, which asked an almost identical question. Reanalyzing the responses to this question, Brick *et al.* (1997) report that 39 percent believed Census records were available to other agencies, 9 percent believed they were not, and 51 percent said they did not know. These figures are quite similar to those obtained in 1995, although a larger percentage answered Don’t Know and a smaller percentage offered the incorrect response in 1979.

Near the end of the 1996 interview, respondents were asked for the first time whether the Census Bureau was forbidden by law from sharing identified data with other agencies, or (in a

---

<sup>4</sup> When respondents who answered Not Sure were asked to guess, the proportion giving the correct response increased to 52.8 percent and 60.5 percent in 1999 and 2000, respectively, for the question about confidentiality, and to 20.8 percent and 24.2 percent for the question about other agencies (calculated from Table 1 in Tourangeau, Singer, and Presser, forthcoming).

split-ballot version) whether the Census Bureau was required by law to keep the data confidential. These questions were repeated in 1999 and 2000. Trends in responses to the “forbidden by law” question show a large increase in the proportion giving the correct response (“Forbidden by law”) between 1996 and 1999, and a further proportional increase between 1999 and 2000 (Singer, *et al.*, 2001, Table 2.18), although even in 2000 the correct response was given by less than half the sample. Incorrect responses also increased between 1996 and 1999, but this trend was dramatically reversed in 2000, perhaps as a result of the Census Bureau’s public relations campaign in connection with the decennial census. In every year, the proportion believing that there is a law requiring confidentiality is much larger than the proportion believing that there is a law forbidding data sharing with other agencies (*ibid.*, Tables 2.18 and 2.19).

However, just as the percentage of those correctly perceiving the Census Bureau’s protection of confidentiality increased between 1995 and 2000, so did the percentage of those saying it would bother them “a lot” if another government agency got their answers to the census, along with their name and address, or if their answers to the census were not kept confidential. The percentage responding “a lot” to the former question increased significantly from 36.8 percent to 45.6 percent between 1995 and 2000 (*ibid.*, Table 2.16); corresponding responses to the latter question increased from 36.6 percent to 49.6 percent between 1996, the first time the question was asked, and 2000 (*ibid.*, Table 2.17). In both cases, the largest increase occurred between 1996 and 1999, with the further change between 1999 and 2000 not statistically significant.

In all three years, respondents who indicated that there were laws forbidding data sharing or requiring confidentiality were asked whether they trusted the Census Bureau to obey these laws. In all three years, about two thirds of those asked said they would trust the Census Bureau to uphold the law (*ibid.*, Table 2.20).

Almost at the end of the questionnaire, respondents were asked three questions designed to measure the prevalence of suspicions sometimes voiced about the misuse of census data for law enforcement purposes. The first of these asked (in 1995, 1999, and 2000), “Do you believe the police and the FBI use the census to keep track of troublemakers?” The percentage of those giving the correct response (i.e., that it is not used for that purpose) increased slightly, from 49.0 percent to 52.1 percent, between 1995 and 1999, and then substantially, to 63.5 percent, between 1999 and 2000 (the overall change is statistically significant at .001). The second question, asked only in 1999 and 2000, asked, “How about to locate illegal aliens? Do you believe the census is used for that?” The percentage voicing this belief declined significantly, from 50.3 percent in 1999 to 42.1 percent in 2000. Finally, respondents in 1999 and 2000 were asked, “Do you agree or disagree that people’s answers to the census can be used against them?” The percentage agreeing declined from 39.2 percent to 37.3 percent, but this change was not statistically significant (*ibid.*, p. 35).

Martin (2001) reports that agreement that people’s answers can be used against them increased significantly during the period following mailing of the Census 2000 forms; her finding is based on Inter-Survey tracking surveys with independent samples. The Surveys of Privacy Attitudes (SPA) also found an increase in agreement with this statement by interview date (logged) during the 2000 survey ( $p=.2$ ) (Singer, 2001, footnote 9). Thus, there is evidence of a long-term decline in suspicion about the use of census data for law enforcement purposes, coupled with



heightened suspicion during the data collection period itself.

## **2.4 Trends in attitudes toward privacy**

So far, we have considered trends in beliefs about confidentiality. SPA also, however, asked questions about privacy, as distinct from confidentiality. One question asked specifically whether the respondent regarded the Census Bureau's asking about age, race, and sex, along with name and address, as an invasion of privacy; others were more general questions. Some of these questions were asked in all four years; most were asked only in 1995, and then again in 1999 and 2000.

There was a small but significant decline between 1995 and 2000 in the percentage of the sample regarding the questions asked on the census short form as an invasion of privacy; 23.5 percent regarded it as an invasion in 1995, and 20.9 percent did so in 2000 (Singer *et al.*, 2001, Table 2.6). Scores on the Privacy Index, consisting of answers to the five more general privacy questions, also declined slightly but significantly during this period.

## **2.5 Attitudes toward confidentiality in the Puerto Rican sample**

A small sample of Puerto Rican residents was interviewed by telephone as part of the SPA in 2000. This sample expressed less concern about privacy and more trust in the Census Bureau to protect confidentiality than the mainland sample did (Singer *et al.*, pp. 107). It is possible that nonrespondents to the survey--the 43 percent of the sample who did not answer the survey--and those (roughly one third of the population) who do not own a telephone may hold attitudes quite different from those of the interviewed sample. It was not possible to match the census returns of the Puerto Rican sample to their survey responses in order to examine the relation between privacy attitudes and census returns directly. Focus group discussions of why Puerto Ricans did not mail back their census forms revealed concerns about confidentiality as one factor mentioned by participants (Berkowitz and Brudvig, 2001, pp. 17-18), but they cannot inform us about the importance of that factor relative to others.

## **2.6 Trends in attitudes toward data sharing**

Singer, Schaeffer, and Raghunathan (1997) have shown that opinions about data sharing are related in predictable ways to trust in government, to confidence in the Census Bureau's promise of confidentiality, to feelings of political effectiveness, and to a more general inclination to share or withhold personal information. Although such opinions may shift in response to media attention to the issue (Kerwin and Forsyth, 1998, p. 19), they can usefully be regarded as reflecting these general predispositions.

Trends in attitudes toward three different issues are explored in the series of surveys under discussion here: the use of administrative records to reduce the undercount; the use of such records to replace the conventional census; and the use of administrative records to provide the information currently collected by means of the census long form.<sup>5</sup> Questions about reducing the

---

<sup>5</sup> The findings of earlier studies are reviewed in Blair (1994).

undercount were asked in terms of data sharing by the Social Security Administration, the Internal Revenue Service (IRS), and one additional agency, which varied from year to year;<sup>6</sup> the order in which agencies were asked about was randomly rotated. Questions about a records-only census did not specify any particular agency, and questions about the long form were asked only about the IRS and a second agency, which also varied from year to year. In each case, the question about administrative records was preceded by a short description of the problem their use was designed to address. Thus, respondents were first informed about the existence of the undercount, and then asked how they felt about specific federal agencies sharing data with the Census Bureau in order to “identify people who are missed in the census.” To motivate the use of administrative records to replace the conventional count, respondents were told, “No one would be asked to fill out a [census] form. Instead, the Census Bureau would count the entire population by getting information from other government agencies.” The question about replacing the long form was preceded by a question probing awareness of the existence of the long form, and the question itself contained a fairly lengthy rationale: “Other government agencies . . . already have some of the information asked on the long form. It has been proposed that they give this information to the Census Bureau. Combining information from agencies would mean that everyone could fill out the short form instead of some people having to fill out the longer form. To make this possible, would you favor or oppose . . .”

All three questions show a decline in approval for data sharing between 1995 and 2000, and in every year, those strongly opposed outnumber those strongly in favor (Singer *et al.*, 2001, Tables 2.21-2.28). The decline in support for a records-only census is almost linear, with the total drop between 1995 and 2000 amounting to approximately seventeen percentage points (*ibid.*, Table 2.29). Those opposed to a records-only census were then asked whether they would favor it if it led to increased accuracy and (if they were still opposed) if it saved money.<sup>7</sup> The argument about accuracy persuaded more people than the argument about economy (*ibid.*, Tables 2.30 and 2.31). The percentage remaining opposed in the face of both arguments increased from 16 percent in 1996 to 23 percent in 1999 and 24 percent in 2000 (*ibid.*, pp. 43-44). Those who remained opposed were asked about the reasons for their opposition. The most frequently given reasons involved concerns about privacy and confidentiality, given by 22 percent in 1999 and 29 percent in 2000; the second most frequent reason was a belief that such a census would be less accurate (17 percent in 1999 and 19 percent in 2000) (*ibid.*, p. 44).

Since 1995, the Surveys of Privacy Attitudes have inquired whether people were aware of the long form and whether they would be willing to have government agencies share data with the Census Bureau in order to eliminate it. Only about one fifth of the population was aware of the existence of the long form in 1996, down somewhat from 1995, and that figure had declined to

---

<sup>6</sup> In 1995, it was the Immigration and Naturalization Service; in 1996, it was the Food Stamp Office; and in 1999-2000 it was “agencies providing public housing assistance.”

<sup>7</sup> The order of asking about accuracy and economy was randomized, with those who continued their opposition to a records-only census being offered the second reason for changing their views.

17 percent in 1999. But by the time of the 2000 survey, which went into the field the week after census forms were delivered to every U.S. household, some 59 percent claimed awareness of the long form (*ibid.*, Table 2.32). However, increased awareness did not translate into increased approval of having government agencies such as the IRS share data with the Census Bureau in order to eliminate the need for the long form. The percentage favoring data sharing for this purpose declined from 52.2 percent in 1995 to 42.9 percent in 2000, at an average rate of about two percentage points per year (*ibid.*, Table 2.33); and, as in the case of data sharing to reduce the undercount, those strongly opposed to data sharing of long-form information outnumbered those strongly in favor by roughly two to one (*ibid.*, Tables 2.34 and 2.35).

In every year, the public was more reluctant to permit sharing of sensitive data than to permit sharing of the information needed to produce a count of the population. However, the gap between the long and short form actually declined, over the years, from about 18 percentage points in 1995 to about 12 percentage points in 2000, because reluctance to permit sharing even short-form information by the IRS declined at a greater rate (*ibid.*, Table 2.36). Not unexpectedly, those who believed the Census Bureau protects data confidentiality were significantly more willing to have other agencies share long-form data with the Census Bureau in all three years (*ibid.*, Table 2.38).

### **3. THE EFFECT OF THE CENSUS OUTREACH CAMPAIGN ON ATTITUDES TOWARD CONFIDENTIALITY AND DATA SHARING**

Chapter 3 of the final report on the SPA assesses the effect of the Census Outreach Campaign on attitudes toward confidentiality and data sharing by comparing attitudes in 1999 and 2000, using the two cross-sectional surveys conducted in those years. The surveys used identical questionnaires, identical methods of sampling and interviewing, had very similar response rates, and were done by the same survey organization, so it is reasonable to assume that they are measuring change in the relevant attitudes. However, the analyses do not control for other demographic characteristics, so it is possible—but very unlikely—that differences in the composition of the sample might account for the differences observed.

The comparison reveals a number of significant changes in attitudes during the ten months separating the two surveys. People's awareness of the uses to which the census is put increased, as did the importance they attached to it. Although there was no change in the percentage—a fifth of the population—who considered the census an invasion of privacy, there was a significant

decline in the percentage of those perceiving, correctly, that other government agencies could not get census data identified by name and address. The percentage of those who knew that the Census Bureau is required by law to protect the confidentiality of the data it collects (or forbidden by law to disclose it) also increased significantly. These changes are, in all likelihood, attributable to publicity about the census, since in most cases they reverse or dramatically accelerate trends apparent from 1995 to 1999 (*ibid.*, p. 3, 59-65).

At the same time, a number of related questions showed no significant change between 1999 and 2000, even though the messages disseminated by government might have been expected to have an impact on responses to at least some of them. First, and perhaps most important, there was no significant increase in the percentage of those who said they believe the government protects the confidentiality of the data. (Given the other findings cited here, we are inclined to interpret the absence of change in responses to this question as signifying that it tapped an element of trust rather than awareness or knowledge about the law.) Nor was there a significant increase in the percentage of those saying they trust the Census Bureau to keep data confidential. This question was asked only of those who answered, correctly, that the Census Bureau is required by law to protect the confidentiality of the data it collects (or prevented by law from disclosing it), a percentage that *did* show a significant increase between the two years. Nor was there any change in the generalized trust which people expressed in the federal government (*ibid.*, p. 3, 59-65).

A series of questions pertaining directly to willingness to have the Census Bureau use data from other agencies to fix the undercount, eliminate the need for a census altogether, or eliminate the need for answering questions on the long form, either showed no change between 1999 and 2000 or, in the case of willingness to have agencies share data to eliminate the census, showed a significant decline. Similarly, willingness to provide one's SSN in order to facilitate such sharing showed no significant change between these two years. It is hard to know how to interpret these findings. A significant decline in willingness to have agencies share data, and to provide one's SSN, had been apparent since 1995 or 1996. This trend appears to have been halted, if not reversed, between 1999 and 2000.

Finally, there does not appear to have been an increase between 1999 and 2000 in concern about either privacy in general or the privacy of census-related information. An indicator of generalized privacy concerns showed a small but significant decline between 1999 and 2000 (*ibid.*, p. 3); however, this decline was no longer significant when demographic variables were controlled (Table 3.1). Responses to the questions that asked, "How much would it bother you if another government agency, outside the Census Bureau, got your name and address, along with your answers to the census?" and "How much would it bother you if your answers to the census were not kept confidential?" showed a similar pattern, with sizable increases in concern from 1995 (or 1996, the first time the question about confidentiality was asked) to 1999, and only small, nonsignificant increases thereafter (*ibid.*, Tables 2.16, 2.17).

Singer *et al.* (2001, pp.66-68) also examined demographic predictors of some of these attitudes. Better-educated respondents in 1999 and 2000 were more knowledgeable about the census and considered it more important than those with less education; they expressed fewer privacy concerns and were less likely to see the census as an invasion of privacy or to believe census information will be misused. They were significantly more likely to believe that other agencies cannot get identified census data and that the Census Bureau protects data confidentiality; they were more willing to have agencies provide data to the Census Bureau to eliminate the long form and to provide their SSN to make this possible.

People whose racial identification was other than white were significantly more concerned about

privacy than whites, less likely to believe the Census Bureau protects confidentiality; less likely to be willing to have agencies share data to reduce the undercount, and less willing to provide their SSN. At the same time, they were more likely to see the census as important than whites.

Self-identification as Hispanic had nonsignificant relationships to many variables, but those that were significant tended to resemble those of the better-educated.

Gender had inconsistent effects on the attitudes measured. Women were less knowledgeable about the census but considered it more important than men do. They were marginally more likely to express trust in government. They were more concerned about privacy in general but less likely to believe that answers to the census would be misused. And though they were significantly more likely than men to favor data sharing under certain circumstances, they were less willing to provide their Social Security number to facilitate this.

The effects of age were also somewhat inconsistent. Older people were significantly more knowledgeable about census uses. They had significantly higher scores than younger people on the general privacy index, but were significantly more likely to believe that other agencies cannot get identified data and less likely to consider the census an invasion of privacy. Nevertheless, they were significantly less likely to trust the Census Bureau to uphold confidentiality laws (and less likely to trust government in general). They were significantly less likely than younger people to approve of any form of data sharing, yet they were significantly more willing to provide their Social Security number to facilitate such sharing.

Like older people, those with higher incomes had significantly higher scores (greater concern) on the privacy index and were significantly less likely than those with lower incomes to trust the Census Bureau to uphold confidentiality laws. Yet they were also significantly less likely than those with lower incomes to think responses to the census would be used against people. Their answers to the data sharing questions are inconsistent.

#### **4. THE IMPACT OF NEGATIVE PUBLICITY ON PRIVACY CONCERNS**

Martin (2000) examined the impact of three factors on privacy and confidentiality concerns during Census 2000: (1) receipt of a long form; (2) mistrust of government; (3) publicity and commentary that increase the public's sensitivity to privacy and confidentiality issues. During Census 2000, several prominent figures commented on the perceived intrusiveness of the long form and were widely quoted in the press in late March. The public became increasingly aware of the controversy, which peaked at just about the time the census forms were being mailed to U.S. households.

Martin (2001) used five cross-sectional surveys conducted by InterSurvey, Inc. under the sponsorship of several private foundations between March 3 and April 13, 2000. The surveys were self-administered using web TV; the cumulative response rate is about 30 percent. Sample sizes ranged from about 1900 for the first survey to 1300 for the fifth. Because of nonresponse and sample biases, the InterSurvey results probably overestimate awareness of the census and

underestimate privacy concerns. However, they probably reflect less bias of the survey auspices, since the Census Bureau was not a sponsor of the surveys.

Martin created a privacy scale consisting of three agree-disagree items: (1) My answers to the census could be used against me; (2) The Census Bureau promise of confidentiality can be trusted; and (3) The census is an invasion of privacy. (The three items form an acceptable unidimensional scale; Martin, 2001, Table 2.) She then tested the effect of three hypothesized causal variables on concern about privacy: (1) receipt of a long form, (2) awareness of the controversy about census long-form questions, and (3) mistrust in the government and general mistrust of people.

Mistrust in government and in people in general was highly predictive of the level of privacy concern about the census (*ibid.*, Table 3), even with demographic variables controlled. Receipt of a long census form was likewise predictive of increased privacy concerns. Finally, hearing about the long form controversy was also strongly associated with increased privacy concerns (*ibid.*). Martin argues, on the basis of several analyses, that this association should be interpreted causally (*ibid.*). The effects of the long form and hearing about the controversy are additive; there is evidence of a weak, nonsignificant interaction between mistrust and hearing about the controversy. These results resemble those reported by Singer *et al.* (2001), who found that respondents who reported exposure to negative as well as positive publicity about the census had significantly higher scores on the privacy index and were significantly more likely to regard the census as an invasion of privacy, and less likely to be willing to provide their Social Security Number, than those reporting no exposure to publicity about the census (*ibid.*, Table 3.2).

Like Singer *et al.* (2001), Martin found fewer privacy concerns among the better-educated and more concerns about privacy among older people. Differences by race and ethnicity were nonsignificant in Martin's data, whereas they were significant in the Singer *et al.* (2001) study (cf. Section 3, above).

Finally, Martin shows that respondents who received a long form or were concerned about privacy were more likely to report returning an incomplete census form, or failing to return it at all (*ibid.*, Table 4). This finding is replicated through actual matching of survey response and census returns in Singer *et al.* (2001); see Section 5, below. An experiment by Junn (2001) also shows that respondents receiving negative priming, in the form of questions designed to raise privacy concerns about the census, were less likely to respond to long-form questions administered experimentally than were respondents who received positive priming, in the form of reasons for asking intrusive questions, or those in a control group, who received no priming at all.

## **5. THE EFFECT OF PRIVACY-CONFIDENTIALITY CONCERNS ON CENSUS PARTICIPATION**

Because the Census Bureau matched the 2000 SPA responses to its file of addresses for the U.S. population, it was possible to examine the relationship between attitudes and behavior--that is,

whether attitudes toward privacy and confidentiality continued to predict census mail returns in Census 2000 as they had in the decennial census a decade earlier (Singer, Mathiowetz, and Couper, 1993). At the conclusion of the interview, all respondents were asked by The Gallup Organization interviewers for their address “in case the Census Bureau wants to do any follow-up research.” (If the address had already been obtained prior to the survey, the interviewer merely verified it with the respondent.<sup>8</sup>) Interviewers obtained 1695 addresses from 1978 respondents, or 85.7 percent.<sup>9</sup> Of these, the Census Bureau matched 1199, or 70.7 percent, at the household level. Thus, analyses of census returns are based on 1199 of the 1978 respondents (60.6 percent) who provided an address that was matched by the Census Bureau (Singer, Van Hoewyk and Neugebauer, forthcoming).<sup>10</sup>

A number of beliefs and attitudes directly related to privacy and confidentiality concerns (Privacy Index; believes census is an invasion of privacy; believes census information may be used for law enforcement purposes), and others inferentially related to these concerns (willing to have agencies share data with Census Bureau to reduce undercount, replace the census, or eliminate the long form; willing to provide SSN), were measured on the 2000 survey. Also measured were positive attitudes toward the census (obligation to cooperate with the census, importance

attached to the census, knowledge about census uses, and trust in the federal government), which the Census Bureau hypothesized would have a positive effect on participation.

Singer, Van Hoewyk, and Neugebauer present three separate tests of the effect of attitudes about privacy and confidentiality on behavior. First, they estimate the effects of these attitudes on willingness to provide an address to the Gallup interviewer. Second, they estimate the effect of privacy and confidentiality concerns on respondents’ return of their census form, correcting these estimates for attrition (due to failure to provide an address and failure to match the address)

---

<sup>8</sup> Where possible, Gallup matched the sample of telephone numbers to lists of addresses before the survey in order to send advance letters to potential respondent households.

<sup>9</sup> The results reported here differ somewhat from those in the final report to the Census Bureau because only the data for 2002 are analyzed here. For a variety of reasons, the prediction from 1999 attitudes to 2000 behavior was deemed unreliable.

<sup>10</sup> This is a very low match rate. For example, for their analysis of privacy and confidentiality as factors in response to the 1990 census, Singer, Mathiowetz and Couper (1993) used respondents to the Survey of Census Participation, carried out in the summer of 1990 by the National Opinion Research Center (NORC) with a response rate of 89.8 percent. Respondents to this survey had been linked to decennial census information as part of a larger project on survey participation (see Groves and Couper, 1998); 97.6 percent of the addresses were successfully matched at the household level. Since the Survey of Census Participation was a face-to-face survey, good addresses were available for all or almost all respondents.

in order to compensate to some extent for the low match rate. Third, to avoid some ambiguities in this analysis, they repeat it for one-person households. Only the second and third are discussed here.

Among the attitudinal variables, the belief that the census may be misused for law enforcement purposes, as measured by an index based on responses to three separate questions, was a significant negative predictor of returning the census form. Thus, like Singer, Mathiowetz, and Couper (1993), the authors found that concerns about the possibility of confidentiality breaches are negative predictors of cooperation with the census. Those who favored data sharing to permit replacing the traditional census with administrative records were also marginally less likely to return their census form. On the other hand, agreeing that everyone has an obligation to cooperate with the census had a significant positive effect on census returns. Concerns about privacy and confidentiality were estimated to explain 1.19 percent of the variance in census mail returns, compared with 1.3 percent in 1990. Thus, the effect of these concerns on mail returns is apparently consistent from 1990 to 2000.

Another way of looking at the effect of confidentiality concerns is to look at the relationship between beliefs that the census may be misused for law enforcement purposes and return of the census form by mail. Of the 478 respondents (39.6 percent of the matched sample) who believed that census data are used for none of the three purposes (identifying illegal aliens, keeping track of troublemakers, and using census answers against respondents), 86.2 percent returned their census form by mail. The percentage dropped to 80.5 percent among those who endorsed one of the three items (N=303), to 76.1 percent among those who endorsed two items (N=255), and to 73.7 percent among the 171 respondents who endorsed all three items. In 1990, census return rates declined from 77.8 percent to 54.8 percent on a similar, but not identical, index of confidentiality concerns. The generally higher return rates in 2000 reflect the fact that for a variety of reasons, the 2000 sample included more compliant respondents than the 1990 sample did. Nevertheless, concerns about confidentiality affect behavior in both samples. The effect in 2000 appears to be linear, whereas the effect in 1990 appeared only among those below the midpoint on the confidentiality index.

The results reported in the paragraph above are bivariate relationships. Controlling for all the other variables included in the multivariate model estimated by Singer, Van Hoewyk, and Neugebauer reduces the percentage spread by a half a percentage point. Given the cost of retrieving census information not returned by mail, even this reduction in the likelihood of returning the census form is substantial.

There are two sources of ambiguity in the data about the effect of privacy and confidentiality concerns on census mail returns. One is that a household may have moved between Census Day (April 1) and the date of the interview, which ranged from a few days to three months later, and this may have attenuated the relationship between attitudes and behavior as measured in the 2000



study.<sup>11</sup> An additional source of ambiguity is that except in one-person households, the individual whose attitudes were measured on the survey was not necessarily the same person who had returned the census form. Respondents were selected at random from household members, and were not asked whether they had returned the census form a few days or weeks earlier.

In order to try to remove this second source of ambiguity, the authors examined the relationship between attitudes and behavior separately in one-person households, where the respondent and the person returning or failing to return the form would almost always be the same. Comparing the results with those in multiple-person households, they found, as expected, that attitudes of all kinds accounted for a much larger share of variance in one-person than in multiple-person households--7.4 percent, compared with 3.1 percent. Privacy attitudes accounted for 4.5 percent of the variance in one-person households, compared with only 1.1 percent in multiple-person households. Thus, the authors argue that the findings concerning the effect of privacy and confidentiality concerns on return of the census form would have been even stronger had they been able in all cases to interview the person who actually returned (or was responsible for returning) the census form for the household.

## **6. PUBLIC PERCEPTIONS OF AGENCY CONFIDENTIALITY PRACTICES**

There is abundant evidence, dating at least to 1979, that public perceptions of agency practices with respect to confidentiality are inaccurate (see Section 1 of this report). In the first place, most people, when given an opportunity, claim not to know what Census Bureau practices with respect to confidentiality are. Second, among those who say they do know, the majority believe that other agencies can gain access to the data. Although there was increased accuracy in public perceptions about agency practices between 1999 and 2000, it is not at all clear that these gains in accuracy will be maintained once the effects of the extensive and intensive public relations campaign connected to the decennial census have faded.

These perceptions of government agency practices are also exemplified in “Privacy Schemas and Data Collection: An Ethnographic Account,” carried out by Gerber as part of the Census 2000 Testing and Experimentation Program. Gerber and her colleagues interviewed 120 people from diverse backgrounds in the DC area, using semi-structured research protocols (*ibid.*, p. ii). They

---

<sup>11</sup> From March 1999 to March 2000, some 16.9 percent of U.S. households moved (U.S. Census Bureau, Current Population Survey, March 2000); therefore some 5 percent of the households interviewed in 2000 may have been occupied by residents other than those who filled out the census form. The respondents were not asked whether they had lived at the same address on Census Day. In the 1990 Survey of Census Participation (SCP), some 6 percent of the sample had moved between Census Day and the date on which they were interviewed (Singer, Mathiowetz, and Couper, 1993). Interviewing for our survey began immediately after Census Day, whereas interviewing for the SCP did not begin until July; as a result, the percentage of movers in the current study should have been even smaller.

found that respondents (many of them members of minority groups) believed that information is freely shared between government agencies, despite assurances of confidentiality. As a result, if they have something to hide, they are reluctant to provide information to any government agency (*ibid.*, pp. iii, 12-14).

## **7. THE EFFECT OF A REQUEST FOR SOCIAL SECURITY NUMBERS**

For more than a dozen years, the Census Bureau has been exploring the use of administrative records in order to improve coverage in the decennial census or to reduce the burden of responding by obtaining some information from records that otherwise would be requested on the census long form. In Census 2000, the Census Bureau sponsored research designed to assess the effect of a request for Social Security Numbers (SSNs) on (a) return rate to the census, (b) item nonresponse, and (c) accuracy of response. In addition, the experiment was designed (d) to assess the effect of a general vs. a specific notification that administrative records might be linked with census records. “General” notification informed the household that census data might be linked with data from other federal agencies, whereas “specific” notification named the agencies whose data might be sought for linking. The experiment was carried out in the context of Census 2000.

Past research had indicated two different response patterns to SSN requests. When asked in the context of a focus group, large majorities react negatively to the prospect of such a request (Singer *et al.*, 1992). In response to a hypothetical question on a survey, a substantial and growing minority indicate they would oppose such a request (Singer, 2001). But when respondents were actually asked for their SSN in a 1992 experiment, the request generated a much smaller than expected (3.4 percent) decline in response rates, and an additional 17 or so percentage point increase in item nonresponse (Dillman, Sinclair, and Clark, 1993). There had been no prior experimental research on notification of administrative record use, and qualitative research had yielded only ambiguous clues (Aguirre International, 1995).

The Census 2000 Social Security Number, Privacy Attitudes, and Notification (SPAN) experiment was designed to clarify all these questions. The experiment created ten panels, three using the long form and seven the short form, with half the forms mailed to High Coverage and half to Low Coverage Areas (Guarino, Hill, and Woltman, 2001, p. 4). The mailout for each panel was a little over 5200; about 10 percent were undeliverable and were excluded from the denominator in calculating response rates (*ibid.*, pp. 6-7).

The findings from the SPAN experiment are consistent with earlier research on actual SSN requests. Asking for a SSN for one or all members of the household results in a small, significant decrease in mail response to Census 2000. When the request is for all household members, it results in a decline of 2.1 percent in High Coverage Areas (HCAs) and 2.7 in Low Coverage Areas (LCAs) (*ibid.*, p. 17). These two figures are not statistically different from each other, and are comparable to the 3.4 percent decline observed in 1992. (Low Coverage Areas contain a large proportion of the country’s Black and Hispanic populations as well as renter-occupied housing units.) The HCA stratum comprised about 81 percent of the total Decennial

Master Address File as of September 9, 1999 (*ibid.*, p. 5).

The findings with respect to notification of administrative record use indicate that such notification results in an additional small but significant decrease in response rates (*ibid.*, Table 3.). Looking at the interaction between notification and the request for SSN suggests that specific notification coupled with such a request results in a significant decline in return rates, whereas a general notification coupled with such a request does not (*ibid.*).

Item nonresponse was defined as the likelihood of a household having any missing data among the 100 percent person items in addition to household tenure. Thus, this data quality measure does not specifically address the effect of an SSN request on providing the SSN itself. The request for an SSN for Person 1 did not result in a significant increase in item nonresponse on the census form, whereas the request for SSNs for all household members did result in such an increase. Notification in the presence of an SSN request did not further increase item nonresponse (*ibid.*, Table 4).

With respect specifically to the SSN item, the results are highly comparable to those by Bates (1992). Some 15.5 percent of SSNs are missing for Person 1 when a request is made for Person 1 only, with increasing percentages missing for Persons 2 through 6 when SSNs are requested for all members of the household (Guarino, Hill, and Woltman, 2001, Table 5). The implication is that those numbers are missing not only because of refusal, but also because the person filling out the census form is ignorant of the SSNs of other household members. Interestingly enough, nonresponse to the SSN item *decreased* for Person 1 in the presence of specific or general notification; such notification had little, if any, effect on nonresponse to the requested SSNs for other household members (*ibid.*, p. 22 and Table 7). This finding supports the interpretation that nonresponse to the request for one's own SSN signifies refusal, whereas nonresponse to the request for others' SSN signifies both refusal and ignorance. It further suggests that the notification statement may have provided justification for providing the SSN number requested. *Thus, whereas specific notification in the presence of an SSN request discouraged return of the census form, specific as well as general notification encouraged supplying the SSN among those who did return the form.*

The Census Bureau subsequently validated the SSNs given by comparing them to the Census Numerical Identification (Numident) File.<sup>12</sup> The results of that exercise showed that 94.77 percent of all those SSNs given were accurate (Brudvig, 2003, p. iv). There was a small but statistically significant difference between low and high coverage areas—95.15 percent were accurate in HCA areas, compared with 92.8 percent in LCA areas. For Person 1, the accuracy rate is even higher, ranging from 96.01 to 96.93 percent depending on the experimental condition. Neither the type of notification, nor requesting an SSN for Person 1 only or for all persons in the household, affects the accuracy of the report for Person 1. However, accuracy

---

<sup>12</sup> Accurate SSNs are those where the SSN and name provided by the respondent match an SSN, name and, as needed, year of birth and gender on the Census Numident file. See Brudvig, 2003, pp. 8-9 for details.

shows a decline in each panel from Person 2 through Person 6. The validation rates found in the current study are very similar to those found in the Simplified Questionnaire Test in 1992 (*ibid.*, pp. iv-v).

The Census Bureau had planned various other analyses of the costs and benefits of using the SSN to link administrative and census files. These analyses were not carried out because of resource constraints. We return to consideration of these analyses in the section on recommendations.

## **8. THE ROLE OF THE PARTNERSHIP PROGRAM IN PRIVACY ATTITUDES**

The Survey of Partners (Westat and Poyer, 2002) asked Partners in the Census how helpful the Partnership public relations program materials provided by the Census Bureau were in achieving each of six stated goals. The ratings showed that materials related to goals of basic education about the census--understanding the purpose of Census 2000 and explaining its importance--received highest ratings (*ibid.*, p. 26). And indeed, the Surveys of Privacy Attitudes showed significant gains in knowledge and significant increases in importance attached to the census between 1999 and 2000 (Singer *et al.*, 2001, pp. 59-60). Materials were rated as less helpful for goals that were related to attitudes of trust in the promise of confidentiality and to creating a sense of civic responsibility for being counted (Westat and Poyer, p. 26). Indeed, materials related to the goal of instilling trust in the Census Bureau's promise of confidentiality were most likely to be rated as not helpful (13 percent), least likely to be rated very helpful (29 percent), and received the highest percentage of Don't Know responses (31 percent) (*ibid.*, p. 26). The Surveys of Privacy Attitudes showed no significant gain in trust between 1999 and 2000, when the Census outreach program was in full swing (Singer *et al.*, 2001, p. 62). These findings suggest that increasing trust in the Census Bureau is likely to be a far more difficult challenge than simply increasing knowledge about the census, but it is necessary if cooperation with the census is to be improved.

## **9. ATTITUDES AND BEHAVIOR**

Evidence on the relevance of the privacy attitude-behavior relationship for the Census Bureau comes from two sources. In the first place, the analyses reported earlier (see Section 5) of the effects of privacy attitudes on return of the census form indicate quite clearly that those who express more concern about privacy and confidentiality are less likely to provide their address and less likely to return their census form, but this relationship is quite small, accounting, in the case of census returns, for less than 2 percent of the variance.

The second test of the relation between attitudes and behavior comes from comparison of answers expressed in various venues toward the hypothetical question of whether respondents would be willing to provide their SSN to the Census Bureau, and the actual provision of those numbers in the SPAN experiment.

Unlike the examination of privacy/confidentiality attitudes and census returns, attitudes and

behaviors with respect to SSNs were measured on different samples of individuals. Hence, there is no question of predicting behavior from attitudes. However, the object of the attitude and the behavior is virtually identical--i.e., willingness to provide one's SSN to facilitate Census Bureau use of other government records. Hence, it is pertinent to ask whether the attitudinal and behavioral measures of willingness to provide one's SSN to the Census Bureau would lead one to the same conclusion.

Singer *et al.* (2001) report that the percentage of those who said they were willing to provide their SSN to the Census Bureau declined from 68.3 percent in 1996 to 55.9 percent in 2000 (Table 2.45). In the SPAN experiment (see Section 7, this report), approximately 3 percent failed to return their census form because of a request for SSN and an additional 15.5 percent failed to provide their SSN for Person 1 if they did return the form. (Nonresponse to the SSN request increased for Persons 2-6 on the form, suggesting that ignorance as well as unwillingness is a factor for these persons.) In addition, between 5 and 7 percent of SSN numbers for Person 1 were inaccurate (Section 7). Thus, this direct test of how closely expressed unwillingness to provide one's SSN corresponds with the number who fail to provide their SSN when asked to do so suggests that approximately half of those saying they would be unwilling to provide their SSN to the Census Bureau would actually fail to provide an accurate number if they were directly asked to do so. This is in fact a very substantial relationship between an attitude expressed in one context and behavior observed in another, and the differences between the two contexts are such as to make the increase in observed compliance plausible. It should be noted once again that this comparison does not involve the same individuals. However, it does involve a comparison between two samples drawn from the same population at about the same point in time.

The question has been raised whether asking everyone for an SSN in the context of the decennial census would facilitate response or increase resistance. This question cannot be answered in the abstract. Experience with Census 2000 suggests that if the request comes to the attention of influential elites who oppose it, compliance might indeed suffer. This was the case with questions on the long form in Census 2000, which generated widely reported unfavorable publicity in the days immediately surrounding distribution of the Census 2000 forms. Analysis of those reporting exposure to this unfavorable publicity suggests that there was indeed a significant negative impact on attitudes (Singer *et al.*, 2001, pp. 69-74; Martin, 2001), including a significant negative effect on expressed willingness to provide one's SSN. However, because of the nonexperimental design, selective exposure is an alternative explanation of the findings.

## **10. THE ROLE OF INFORMED CONSENT**

The question is often raised whether providing respondents information about the content of a survey ahead of time is likely to increase or reduce their willingness to participate in it. The question cannot, of course, be answered in that form. For example, while an experiment by Singer (1978) shows that providing respondents somewhat more information about the sensitive content of a survey did not reduce the response rate or increase refusals, an experiment by Groves, Presser, and Dipko (2000) indicates that respondents who, by virtue of their membership in certain organizations, might be expected to be more interested in a particular topic were in fact more likely to respond when the introduction mentioned that topic. It is also possible that

respondents who hold very negative attitudes about a particular topic may refuse to participate in a survey about that topic, or they may be sensitized by a reference to the topic in the introduction to refuse to answer certain questions (Singer, 1978).

The Social Security Notification experiment indicates that informing respondents that their census form might be linked to other government records had a small but significant negative effect on returns of the census form itself (see Section 6 above). However, among those who did return the form, notification had a positive effect on the likelihood of responding to the Social Security Number question (Section 6). Notification had no impact on whether the SSN reported was accurate. Thus, as an empirical matter, the value of informing respondents about linkage ahead of time may depend on whether the interest is primarily in return of the entire form or in answers to the question about SSN. As an ethical matter, however, some would argue that there is an obligation to inform respondents about the proposed linkage in any case.

In her ethnographic exploration of the meaning of privacy and confidentiality concerns and the relevance of these concerns for returning the census form, Gerber (2002) points out that respondents form expectations of what questions are legitimate for a sponsor to ask, based on their understanding of the nature and purpose of the survey and the sponsoring organization (*ibid.*, p. iii). She therefore recommends that sponsors provide good, understandable explanations of why these data are needed and how they will be used (*ibid.*, p. iv). This recommendation would clearly seem to apply to the Census Bureau's request for SSNs. Martin's ASA paper (2001) also addresses these issues, and includes information about the nature of respondents' objections to long form questions and about the kind of information they say they want.

## **11. SUMMARY OF FINDINGS**

### **11.1 Trends in attitudes, 1995-2000**

In general, knowledge about Census Bureau confidentiality practices increased between 1995 and 2000. Even in 2000, however, most respondents either did not know what the Census Bureau's practices and legal obligations were, or responded that data were shared with others.

During this same period of time, the percentage saying they would be bothered "a lot" if their census data were shared with anyone outside the Census Bureau increased significantly, and approval of data sharing for all three of the purposes asked about (to reduce the undercount, to eliminate the census, and to replace the long form) declined. With the exception of support for data sharing to reduce the undercount, which stabilized between 1999 and 2000, these trends appear to be linear. Expressed willingness to provide one's Social Security number also declined between 1996 and 1999, with no further change in 2000.

Increased disapproval of data sharing was not paralleled either by increasing distrust of the uses to which census data might be put, or by increasing concerns about privacy in general, or by declining trust in government or in the Census Bureau to keep data confidential. Three questions about possible misuse of census data all showed a decline in distrust between 1999 and 2000,

with two of the three statistically significant. The question asking whether people trust the Census Bureau to keep data confidential (if they correctly perceived that there were laws governing confidentiality) showed no significant change. The question asking whether the census short form is an invasion of privacy showed a small significant decline between 1995 and 2000, and other questions asking about general privacy concerns for the most part showed no consistent trends. Finally, people's trust in "the government in Washington" showed a small, significant increase between 1996 and 2000 after declining from 1995 to 1996.

### **11.2 The effect of Census Bureau public relations activities on beliefs about confidentiality**

Although there was no change between 1999 and 2000 in the percentage--a fifth of the population--who considered the census an invasion of privacy, there was a significant decline in the belief that census data were likely to be misused, and a significant increase in the percentage of those perceiving, correctly, that other government agencies could not get census data identified by name and address. The percentage of those who knew that the Census Bureau is required by law to protect the confidentiality of the data it collects (or forbidden by law to disclose it) also increased significantly. These changes are, in all likelihood, attributable to publicity about the census commissioned or stimulated by the Census Bureau, since in most cases they reverse or dramatically accelerate trends apparent from 1995 to 1999.

At the same time, a number of related questions showed no significant change between 1999 and 2000, even though the messages disseminated by the Census Bureau might have been expected to have an impact on responses to at least some of them. First, and perhaps most important, there was no significant increase in the percentage of those who said they believe the government protects the confidentiality of the data. (Given the other findings cited here, we are inclined to interpret the absence of change in responses to this question as signifying that it tapped an element of trust rather than awareness or knowledge about the law.) Nor was there a significant increase in the percentage of those saying they trust the Census Bureau to keep data confidential.

### **11.3 Attitudes and behavior**

As we have seen, the relationship between attitudes and behavior varies depending on how close the conceptual relationship is between the two. Respondents' concerns about confidentiality and privacy predict their return of the census form. The relationship to census returns is statistically significant, and though small, it is of the same order of magnitude as that measured in the 1990 census (Singer, Mathiowetz, and Couper, 1993). Obviously, there is no one-for-one relationship between concerns about confidentiality and willingness to return the census form; the latter is motivated by other attitudes and concerns -- e.g., how important the respondent considers the census, concern about breaking the law, desire to obtain a fair share of government resources--and these attenuate the relationship between privacy/confidentiality concerns and census returns. The relationship is also attenuated by error in the measurement of both attitudes and behavior. Looked at another way, however, the relationship is substantial. The difference in census returns between those who endorse none of the three beliefs that the census may be

misused for law enforcement purposes and those who endorse all three of those beliefs is 12.5 percentage points.

When the object of the attitude and the behavior are more closely related to one another, the conclusions that would be drawn from attitudinal and behavioral data are easier to compare. Some 45 percent of respondents to the 1999 and 2000 surveys said they would not be willing to provide their SSN to the Census Bureau to facilitate data linkage; in an experimental test of such willingness, some 22 percent of subjects failed to provide an accurate SSN for Person 1 on the census form. Thus, one might conclude that about half of those expressing reservations about providing a Social Security Number would actually fail to do so if requested by the Census Bureau.

#### **11.4 Willingness to provide SSNs**

The SPAN experiment demonstrates that (a) large percentages of the population will provide their SSN to the Census Bureau if asked to do so on the census form; (b) the request reduces the response rate by less than 3 percentage points; (c) nonresponse to the SSN item totals 15.5 percent for Person 1 and more for subsequent persons in the household; (d) between 93 and 95 percent of the SSNs provided are accurate; and (e) notification of record linkage has a small but significant negative effect on the response rate but a small positive effect on responding to the SSN item.

## **12. CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE RESEARCH**

### **12.1 Conclusions**

In this section, I outline the conclusions that can be drawn on the basis of the research carried out under the Census Bureau's privacy initiative, keeping in mind the caveats outlined at the beginning of this report.

1. Except in the period surrounding the decennial census, when publicity about the census was at its height, knowledge and beliefs about the Census Bureau and attitudes toward privacy and confidentiality showed only small year-to-year changes between 1995 and 2000.

Nevertheless, there appears to be a trend toward increasing concern about the sharing of confidential data.

2. Knowledge, beliefs, and attitudes were all significantly related to self-reported exposure to positive as well as negative publicity about the census. However, the SPA found no direct effect of self-reported exposure on census returns. Knowledge and opinions seemed much more amenable to change as a result of the publicity campaign than trust.

3. Attitudes about privacy and confidentiality vary by demographic characteristics, with



nonwhites significantly more concerned about privacy than whites, less likely to believe the Census Bureau protects confidentiality; less likely to be willing to have agencies share data; and less willing to provide their SSN.

4. Attitudes about confidentiality and privacy are reliable and significant predictors of behavior. They predicted respondents' return of the census form in 2000 as well as in 1990, with concern about the possible misuse of census data for law enforcement purposes accounting for a substantial decline in census mail returns.

5. Approval of data sharing among federal agencies, as well as expressed willingness to provide one's Social Security number to facilitate such sharing, declined between 1995 and 2000.

6. At the same time, the impact of a request for SSN on response rates was much smaller than would be predicted from the attitudes expressed, accounting for a decline in response rates of less than 3 percentage points. When nonresponse to the SSN item and inaccuracy in SSN numbers are added to failure to mail back the census form, however, the impact was larger, amounting to some 22 percent of the sample. Thus, whether to request a SSN or not depends on whether the primary goal is to match census returns to other records, or to minimize the cost of the basic enumeration.

7. Ethnographic work in connection with Census 2000 and focus groups in Puerto Rico are in accord with the conclusions reached on the basis of surveys and experiments. Respondents and participants expressed concerns about privacy and confidentiality and mentioned these as reasons for reluctance in sharing information.

8. There is support both from the ethnographic interviews and from the Notification part of the SPAN experiment that explaining why information is requested may help in obtaining it. Even though notification had a small but significant negative impact on response rates, it increased response to the SSN item.

## **12.2 Recommendations**

Given these general conclusions, we make the following recommendations for future research in this area by the Census Bureau:

1. Continue to monitor trends in knowledge, beliefs, and attitudes. Given the small year-to-year changes observed to date, the interval between surveys could probably be increased to three or four years. Such surveys are needed to document the effect of recent events on knowledge and beliefs relevant to the census.

Supplementing the cross-sectional surveys with some panel components would be useful in understanding what causes individual-level change. In addition, it might be useful, and economical, to add a subset of key questions to ongoing high-quality surveys (e.g., the General Social Survey or the Survey of Consumer Attitudes) at more frequent intervals. For both of these activities, replication of questions and survey procedures is essential if the intent is to monitor

change in attitudes.

2. Well before the next census, design, conduct, and analyze small-scale research that develops and then tests more effective ways of communicating the Census Bureau's confidentiality practices to the general public. It is clear from both the ethnographic and survey research reviewed above that the public does not understand or believe the confidentiality assurances provided by the Census Bureau, and that their beliefs have consequences for behavior. A program of qualitative research, followed by small-scale laboratory experiments, is recommended. The effectiveness of the messages developed in this way should then be tested in field experiments.

3. Conduct qualitative research on impediments to trust in the Census Bureau and in the government more generally, and on ways in which feelings of trust might be enhanced. Research of this kind may be especially useful among groups who are less likely to cooperate with the Census Bureau. Because trust, rather than knowledge, may well be crucial to the public's cooperation with the census, this research is especially important. At the same time, such research is likely to be difficult to carry out, and it is not clear that much can be done by the Census Bureau to change levels of trust. Research by economists and psychologists on decision making under conditions of uncertainty, including recent research by neuroscientists, should be reviewed for its possible contribution to the activities in Recommendations 2 and 3.

4. Because attitudes toward privacy and confidentiality account for only a small portion of the variance in census mail returns, design and conduct research to identify and reduce other barriers to response. While it is clear from all the research carried out under this program that concerns about privacy and confidentiality affect respondents' cooperation, it is also clear that such concerns explain only a small part of the variation in behavior. Thus, it is important to look for, and affect, other potential barriers to response, such as motivation and capacity.

## **Acknowledgments**

I would like to thank Randall J. Neugebauer for his advice, help, and encouragement on the 1999 and 2000 Surveys of Privacy Attitudes; Linda Brudvig for commissioning the privacy topic report; and Pat Doyle and her colleagues for inviting me to write a chapter on public perceptions of confidentiality for their book on confidentiality and data access (2001). Thanks are also due to my colleagues at The Gallup Organization and the Survey Research Center at the University of Michigan, in particular Roger Tourangeau and Darby Miller Steiger for help with the design and execution of the surveys and John Van Hoewyk and Trivellore Raghunathan for help with data analysis. In addition to the surveys on privacy attitudes, the excellent research reports by Jennifer Guarino, Joan Hill, and Henry Woltman; by Linda Brudvig; by Susan Berkowitz and Linda Brudvig; by Eleanor Gerber; and by Elizabeth Martin form the basis for this topic report on privacy.

## **References**

Aguirre International. 1995. "Public Concerns About the Use of Administrative Records." Report to the U.S. Census Bureau under contract 50-YABC-2-66027. July.

Bates, N. 1992. "Revised Item Nonresponse Results for Social Security Number from the Simplified Questionnaire Test (SQT)." Memorandum to R. D. Tortora and S.M. Miskura, June 18.

Berkowitz, S. and L. Brudvig. 2001. "Puerto Rico Focus Groups on Why Households Did Not Mail Back the Census 2000 Questionnaire." Washington, D.C., U.S. Census Bureau Planning, Research, and Evaluation Division, July 17.

Blair, J. 1994. "Ancillary Uses of Government Administrative Data on Individuals: Perceptions and Attitudes." Unpublished paper prepared for the National Academy of Sciences Panel on Census Methods, 1994.

Brick, J. J., D. Cantor, J. Kerwin, G. Pamela, and E. Singer. 1997. "Review of Methods to Increase Coverage and Response Rates on the Study of Privacy Attitudes." Report to the U.S. Census Bureau under contract 50-YABC-2-66025.

Brudvig, L. 2003. "Analysis of the Social Security Number Validation Component of the Social Security Number, Privacy Attitudes, and Notification Experiment." Washington D. C.; U.S. Census Bureau; Planning, Research, and Evaluation Division; January 13.

Couper, M.P., E. Singer, and R.A. Kulka. 1998. "Participation in the 1990 Decennial Census." *American Politics Quarterly* 26: 59-80.

Dillman, D.A., M. D. Sinclair, and J. R. Clark. 1993. "Effects of Questionnaire Length, Respondent Friendly Design, and a Difficult Question on Response Rates for Occupant-Addressed Census Mail Surveys." *Public Opinion Quarterly* 57: 289-304.

Fay, R.L., N. Bates, and J. Moore. 1991. "Lower Mail Response in the 1990 Census: A Preliminary Interpretation." In Bureau of the Census, Proceedings of the Annual Research Conference, Washington, DC.

Gerber, E. 2003. "Privacy Schemas and Data Collection: An Ethnographic Account." Washington D.C., U.S. Census Bureau, Statistical Research Division, February 10.

Groves, R. M. and M. P. Couper. 1998. *Nonresponse in Household Surveys*. New York, Wiley.

Groves, R.M., S. Presser, and S. Dipko. 2000. "The Role of Topic Interest in Survey Participation Decisions." College Park, MD, University of Maryland, unpublished paper.

Guarino, J., J. Hill, and H. F. Woltman. 2001. "Analysis of the Social Security Number-

Notification Component of the Social Security Number, Privacy Attitudes, and Notification Experiment.” Washington D.C., U.S. Census Bureau, Planning, Research and Evaluation Division, November 13.

Junn, J. 2001. “The Influence of Negative Political Rhetoric: An Experimental Manipulation of Census 2000 Participation.” Paper presented at the Midwest Political Science Association, April.

Kerwin, J. and B. Forsyth. 1998. "Privacy and Public Attitudes Toward an Administrative Records Census." Report to the U.S. Census Bureau under Contract 50-YABC-2-66025.

Kerwin, J., and S. Edwards. 1996. “The 1996 Study of Public Attitudes Toward Administrative Record Use: Final Report.” Report to U.S. Census Bureau under Contract 50-YABC-2-66025.

Martin, Elizabeth A. 2000. “Public Opinion Changes during Two Censuses.” Paper presented at the Decennial Census Advisory Committee, September 21.

Martin, Elizabeth A. 2001. “Privacy Concerns and the Census Long Form: Some Evidence from Census 2000.” American Statistical Association, *Proceedings of the Section on Survey Research Methods*.

National Research Council. 1979. *Privacy and Confidentiality as Factors in Survey Response*. Washington DC, National Academy of Sciences.

Schuman, H. and S. Presser. 1981. *Questions and Answers in Attitude Surveys*. New York, Academic Press.

Singer, E. 1978. “Informed Consent: Consequences for Response Rate and Response Quality in Social Surveys.” *American Sociological Review* 43:144-62.

Singer, E. 2001. “Public Perceptions of Confidentiality and Attitudes Toward Data Sharing by Federal Agencies.” In P. Doyle, J. I. Lane, J.M.M. Theeuwes, and L.V. Zayatz, *Confidentiality, Disclosure, and Data Access*. Amsterdam, Elsevier.

Singer, E. and E.R. Miller. 1992. "Report on Focus Groups." Washington DC, U.S. Bureau of the Census, Center for Survey Methods Research.

Singer, E., N.A. Mathiowetz, and M.P. Couper. 1993. “The Impact of Privacy and Confidentiality Concerns on Survey Participation: The Case of the 1990 U.S. Census.” *Public Opinion Quarterly* 57: 465-82.

Singer, E., N.C. Schaeffer, and T.E. Raghunathan. 1997. “Public Attitudes Toward Data Sharing By Federal Agencies.” *International Journal of Public Opinion Research* 9: 277-84.

Singer, E., J. Van Hoewyk, R. Tourangeau, D.M. Steiger, M. Montgomery, and R. Montgomery.

2001. "Final Report on the 1999-2000 Surveys of Privacy Attitudes." Washington, D.C., U.S. Bureau of the Census, Planning, Research, and Evaluation Division, December 10.

Singer, E., J. Van Hoewyk, and R. J. Neugebauer. Forthcoming. "Attitudes and Behavior: The Impact of Privacy and Confidentiality Concerns on Participation in the 2000 Census." *Public Opinion Quarterly*.

Steffey, D. and N. Bradburn. 1994. "Counting People in the Information Age." Washington D.C., National Academy Press.

Tourangeau, R., E. Singer, and S. Presser. Forthcoming. "Context Effects in Attitude Surveys: Effects on Remote Items and Impact on Predictive Validity." *Sociological Methods and Research*.

U.S. General Accounting Office. 1992. "Decennial Census: 1990 Results Show Need for Fundamental Reform." GAO/GGD-92-94. Washington, DC, General Accounting Office, 1992.

Westat. 2002. "Report of Survey of Partners." Washington D.C., U.S. Census Bureau, May 21.

Westin, A. 1990. *The Equifax Report on Consumers in the Information Age*. Atlanta, GA: Equifax, Inc.